A CONVERSE COMPARISON THEOREM FOR BSDES AND RELATED PROPERTIES OF $g$–EXPECTATION

PHILIPPE BRIAND
IRMAR, Université Rennes 1
35 042 Rennes Cedex, FRANCE
email: Philippe.Briand@univ-rennes1.fr

FRANÇOIS COQUET
IRMAR, Université Rennes 1
35 042 Rennes Cedex, FRANCE
email: Francois.Coquet@univ-rennes1.fr

YING HU
IRMAR, Université Rennes 1
35 042 Rennes Cedex, FRANCE
email: Ying.Hu@univ-rennes1.fr

JEAN MÉMIN
IRMAR, Université Rennes 1
35 042 Rennes Cedex, FRANCE
email: Jean.Memin@univ-rennes1.fr

SHIGE PENG
Department of Mathematics, Shandong University
Jinan 250100, CHINA
email: peng@public.jn.sd.cn

submitted October 9, 1999 Final version accepted May 23, 2000

AMS 2000 Subject classification: 60H10, 60H30
backward stochastic differential equations, comparison theorem

Abstract

In [1], Z. CHEN proved that, if for each terminal condition $\xi$, the solution of the BSDE associated to the standard parameter $(\xi, g_1)$ is equal at time $t = 0$ to the solution of the BSDE associated to $(\xi, g_2)$ then we must have $g_1 \equiv g_2$. This result yields a natural question: what happens in the case of an inequality in place of an equality? In this paper, we try to investigate this question and we prove some properties of “$g$–expectation”, notion introduced by S. PENG in [8].
1 Introduction

It is by now well-known that there exists a unique, adapted and square integrable, solution to a backward stochastic differential equation (BSDE for short in the remaining of the paper) of type

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s \cdot dW_s, \quad 0 \leq t \leq T,$$

providing, for instance, that the generator is Lipschitz in both variables $y$ and $z$ and that $\xi$ and $(f(s, 0, 0))_{s \in [0, T]}$ are square integrable. We refer of course to E. Pardoux and S. Peng [4, 5] and to N. El Karoui, S. Peng and M.-C. Quenez [2] for a survey of the applications of this theory in finance.

One of the great achievement of the theory of BSDEs is the comparison theorem for real-valued BSDEs due to S. Peng [7] at first and then generalized by several authors, see e.g. N. El Karoui, S. Peng and M.-C. Quenez [2, Theorem 2.2]. It allows to compare the solutions of two BSDEs whenever we can compare the terminal conditions and the generators. In this paper we try to investigate an inverse problem: if we can compare the solutions of two BSDEs (at time $t = 0$) with the same terminal condition, for all terminal conditions, can we compare the generators?

The result of Z. Chen [1] can be read as the first step in solving this problem. Indeed, he proved using the language of “$g$-expectation” introduced by S. Peng in [8] that, given two generators, say $g_1$ and $g_2$, then if, for each $\xi \in L^2$, we have $Y_0^1(\xi) = Y_0^2(\xi)$ where $(Y_t^i(\xi), Z_t^i(\xi))_{t \in [0, T]}$ stands for the solution of the BSDE:

$$Y_t = \xi + \int_t^T g_i(s, Y_s, Z_s)ds - \int_t^T Z_s \cdot dW_s, \quad 0 \leq t \leq T,$$

for $i = 1, 2$, then we have $g_1 \equiv g_2$. With the above notations, the main issue of this paper is to address the following question: if, for each $\xi \in L^2$, we have $Y_0^1(\xi) \leq Y_0^2(\xi)$, do we have $g_1 \leq g_2$?

The paper is organized as follows: in section 2, we introduce some notations and we make our assumptions. In section 3, we prove the result in the case of deterministic generators and we give an application of these techniques to partial differential equations (PDEs for short in the rest of the paper). In section 4, we prove a converse to the comparison theorem for BSDEs and then, with the help of this result we study the case when the generators do not depend on the variable $y$. Finally, in section 5, we discuss the Jensen inequality for “$g$–expectation”.

2 Preliminaries

2.1 Notations and assumptions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space carrying a standard $d$-dimensional Brownian motion, $(W_t)_{t \geq 0}$, starting from $W_0 = 0$, and let $(\mathcal{F}_t)_{t \geq 0}$ be the $\sigma$-algebra generated by $(W_t)_{t \geq 0}$. We do the usual $\mathbb{P}$-augmentation to each $\mathcal{F}_t$ such that $(\mathcal{F}_t)_{t \geq 0}$ is right continuous and complete.

If $z$ belongs to $\mathbb{R}^d$, $\|z\|$ denotes its Euclidean norm. We define the following usual spaces of processes:

- $\mathcal{S}_2 = \left\{ \psi \text{ progressively measurable; } \|\psi\|_{\mathcal{S}_2}^2 := \mathbb{E}\left[ \sup_{0 \leq t \leq T} |\psi_t|^2 \right] < \infty \right\},$
Let us consider a function \( g \), which will be in the following the generator of the BSDE, defined on \( \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \), with values in \( \mathbb{R} \), s.t. the process \( (g(t, y, z))_{t \in [0, T]} \) is progressively measurable for each \((y, z)\) in \( \mathbb{R} \times \mathbb{R}^d \). For the function \( g \), we will use, through-out the paper, the following assumptions.

(A 1). There exists a constant \( K \geq 0 \) s.t. \( \mathbb{P} - a.s., \) we have:
\[
\forall t, \, \forall(y, y'), \, \forall(z, z'), \quad |g(t, y, z) - g(t, y', z')| \leq K(|y - y'| + |z - z'|).
\]

(A 2). The process \( (g(t, 0, 0))_{t \in [0, T]} \) belongs to \( \mathcal{H}_2 \).

(A 3). \( \mathbb{P} - a.s., \forall(t, y), \; g(t, y, 0) = 0. \)

(A 4). \( \mathbb{P} - a.s., \forall(y, z), \; t \mapsto g(t, y, z) \) is continuous.

It is by now well known that under the assumptions (A 1) and (A 2), for any random variable \( \xi \) in \( L^2(\mathcal{F}_T) \), the BSDE
\[
Y_t = \xi + \int_t^T g(s, Y_s, Z_s)ds - \int_t^T Z_s \cdot dW_s, \quad 0 \leq t \leq T, \tag{1}
\]
has a unique adapted solution, say \((Y_t, Z_t)_{t \in [0, T]}\) s.t. \( Z \) is in the space \( \mathcal{H}_2 \). Actually, by classical results in this field, the process \( Y \) belongs to \( \mathcal{S}_2 \).

In [8], S. Peng adopted a new point of view in the study of BSDEs. Indeed, if \( g \) satisfies the assumptions (A 1) and (A 3), he introduced the function \( \mathcal{E}_g \) defined on \( L^2(\mathcal{F}_T) \) with values in \( \mathbb{R} \) by simply setting \( \mathcal{E}_g(\xi) := Y_0 \) where \((Y, Z)\) is the solution of the BSDE (1) (since this solution is adapted, \( Y_0 \) is deterministic). He called \( \mathcal{E}_g \) a g-expectation and he proved that some properties of the classical expectation are preserved (monotonicity for instance) but, since \( g \) is not linear in general, the linearity is not preserved (we will see at the end of the paper that the Jensen inequality does not hold in general for \( \mathcal{E}_g \)).

Related to this "nonlinear expectation", he defined also a conditional g-expectation by setting
\[
\mathcal{E}_g(\xi | \mathcal{F}_t) := Y_t \quad \text{which is the unique random variable } \eta, \mathcal{F}_t \text{-measurable and square-integrable, s.t.}
\]
\[
\forall A \in \mathcal{F}_t, \quad \mathcal{E}_g(1_A \eta) = \mathcal{E}_g(1_A \xi).
\]

Z. Chen in [1] used these notions to prove the following result:

**Theorem 2.1** Let the assumptions (A 1), (A 3) and (A 4) hold for \( g_1 \) and \( g_2 \) and let us assume moreover that, for each \( \xi \in L^2(\mathcal{F}_T), \) \( \mathcal{E}_{g_1}(\xi) = \mathcal{E}_{g_2}(\xi). \) Then, \( \mathbb{P} - a.s., \forall(t, y, z), \; g_1(t, y, z) = g_2(t, y, z). \)

In this note, as mentioned in the introduction, we will try to investigate the following question: if, for each \( \xi \in L^2(\mathcal{F}_T), \) \( \mathcal{E}_{g_1}(\xi) \leq \mathcal{E}_{g_2}(\xi), \) do we have \( g_1(t, y, z) \leq g_2(t, y, z)? \) This problem is roughly speaking a converse to the comparison theorem for BSDEs, since if \( g_1 \leq g_2 \) then \( \mathcal{E}_{g_1}(\cdot) \leq \mathcal{E}_{g_2}(\cdot). \)

We close this subsection by a comment about the assumption (A 3). Under (A 3), if \( \xi \) is a random variable \( \mathcal{F}_S \)-measurable with \( S < T, \) then we have: \( \mathcal{E}_g(\xi | \mathcal{F}_t) = \xi \) if \( S \leq t \leq T \) and
\[ \mathcal{E}_g(x \mid \mathcal{F}_t) = y_t \text{ for } 0 \leq t < S \text{ where } ((y_r, z_r))_{r \in [0, S]} \text{ stands for the solution on } [0, S] \text{ of the BSDE:} \\
\quad y_r = \xi + \int_r^S g(u, y_u, z_u) \, du - \int_r^S z_u \cdot dW_u, \quad 0 \leq r \leq S. \]

This remark will be used in the proofs of Theorem 3.2, Theorem 4.1 and Lemma 4.3.

### 2.2 Technical Results

In this subsection, we establish a technical result which will be useful in the next section. We start by giving an a priori estimate for BSDEs which is of standard type, see [2].

**Proposition 2.2** Let \( \xi \in L^2(\mathcal{F}_T) \) and let the assumptions (A 1) and (A 2) hold. Then, if the process \((Y_t, Z_t)_{t \in [0, T]} \) is the solution of the BSDE (1) we have, for \( \beta = 2(K + K^2) \):

\[
\mathbb{E}\left\{ \sup_{t \leq s \leq T} e^{\beta s} |Y_s|^2 + \int_t^T e^{\beta s} |Z_s|^2 \, ds \mid \mathcal{F}_t \right\} \leq C \mathbb{E}\left\{ e^{\beta T} |\xi|^2 + \left( \int_t^T e^{(\beta/2)s} |g(s, 0, 0)| \, ds \right)^2 \mid \mathcal{F}_t \right\},
\]

where \( C \) is a universal constant.

**Proof.** We outline the proof for the convenience of the reader.

As usual we start with Itô’s formula to see that, noting \( \mathbb{E} = e^{\beta u} \) for \( u \in [0, T] \),

\[
e^{\beta u} |Y_u|^2 + \int_u^T e^{\beta s} |Z_s|^2 \, ds = e^{\beta T} |\xi|^2 + 2 \int_u^T e^{\beta s} Y_s \cdot g(s, Y_s, Z_s) \, ds - \int_u^T \beta e^{\beta s} |Y_s|^2 \, ds - M_u.
\]

Using the Lipschitz assumption on \( g \) and then the inequality \( 2K |\xi| \leq 2K^2 |\xi|^2 + (1/2) |z|^2 \), we deduce that, taking into account the definition of \( \beta \),

\[
e^{\beta u} |Y_u|^2 + \frac{1}{2} \int_u^T e^{\beta s} |Z_s|^2 \, ds \leq e^{\beta T} |\xi|^2 + 2 \int_u^T e^{\beta s} |Y_s| \cdot |g(s, 0, 0)| \, ds - 2 \int_u^T e^{\beta s} Y_s Z_s \cdot dW_s. \tag{2}
\]

In particular, taking the conditional expectation w.r.t. \( \mathcal{F}_t \) of the previous inequality written for \( u = t \), we deduce since the conditional expectation of the stochastic integral vanishes,

\[
\mathbb{E}\left\{ \int_t^T e^{\beta s} |Z_s|^2 \, ds \mid \mathcal{F}_t \right\} \leq 2 \mathbb{E}\left\{ e^{\beta T} |\xi|^2 + 2 \int_t^T e^{\beta s} |Y_s| \cdot |g(s, 0, 0)| \, ds \mid \mathcal{F}_t \right\}. \tag{3}
\]

Moreover, coming back to the inequality (2), we get

\[
\sup_{t \leq u \leq T} e^{\beta u} |Y_u|^2 \leq e^{\beta T} |\xi|^2 + 2 \int_t^T e^{\beta s} |Y_s| \cdot |g(s, 0, 0)| \, ds + 4 \sup_{t \leq u \leq T} \int_t^u e^{\beta s} Y_s Z_s \cdot dW_s,
\]

and by Burkholder-Davis-Gundy’s inequality, the previous estimate yields the inequality

\[
\mathbb{E}\left\{ \sup_{t \leq u \leq T} e^{\beta u} |Y_u|^2 \mid \mathcal{F}_t \right\} \leq \mathbb{E}\left\{ e^{\beta T} |\xi|^2 + 2 \int_t^T e^{\beta s} |Y_s| \cdot |g(s, 0, 0)| \, ds \mid \mathcal{F}_t \right\} + C \mathbb{E}\left\{ \left( \int_t^T e^{2\beta s} |Y_s|^2 |Z_s|^2 \, ds \right)^{1/2} \mid \mathcal{F}_t \right\},
\]

where \( C \) is a universal constant.
where $C$ is a universal constant which will change from line to line; and thanks to the inequality $ab \leq a^2/2 + b^2/2$ we deduce immediately

$$
\mathbb{E}\left\{ \sup_{t \leq u \leq T} e^{\beta u} |Y_u|^2 \bigg| \mathcal{F}_t \right\} \leq \mathbb{E}\left\{ e^{\beta T} |\xi|^2 + 2 \int_t^T e^{\beta s} |Y_s| \cdot |g(s,0,0)| ds \bigg| \mathcal{F}_t \right\} 
+ \frac{C^2}{2} \mathbb{E}\left\{ \int_t^T e^{\beta s} |Z_s|^2 ds \bigg| \mathcal{F}_t \right\} + \frac{1}{2} \mathbb{E}\left\{ \sup_{t \leq u \leq T} e^{\beta u} |Y_u|^2 \bigg| \mathcal{F}_t \right\}.
$$

Combining the inequality (3) with the previous one, we easily derive that, for a universal constant $C$, we have

$$
\mathbb{E}\left\{ \sup_{t \leq u \leq T} e^{\beta u} |Y_u|^2 + \int_t^T e^{\beta s} |Z_s|^2 ds \bigg| \mathcal{F}_t \right\} \leq C \mathbb{E}\left\{ e^{\beta T} |\xi|^2 + \int_t^T e^{\beta s} |Y_s| \cdot |g(s,0,0)| ds \bigg| \mathcal{F}_t \right\}.
$$

It remains to use the fact that

$$
C \mathbb{E}\left\{ \int_t^T e^{\beta s} |Y_s| \cdot |g(s,0,0)| ds \bigg| \mathcal{F}_t \right\} \leq \frac{1}{2} \mathbb{E}\left\{ \sup_{t \leq u \leq T} e^{\beta u} |Y_u|^2 \bigg| \mathcal{F}_t \right\} 
+ \frac{C^2}{2} \mathbb{E}\left\{ \left( \int_t^T e^{(\beta/2)s} |g(s,0,0)| ds \right)^2 \bigg| \mathcal{F}_t \right\}.
$$

and to change $C$ one more time to finish the proof of this proposition. \qed

Before stating our first result, we need some further notations.

(A 5). Let $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n \times d}$ be two Lipschitz functions.

We denote by $X^{t,x}_s$ the solution of the SDE:

$$
X^{t,x}_s = x + \int_t^s b(X^{t,x}_u) \, du + \int_t^s \sigma(X^{t,x}_u) \cdot dW_u, \quad s \geq t,
$$

with the usual convention $X^{t,x}_t = x$ if $s < t$.

Let $g : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be a function s.t. the stochastic process $(g(s,y,z))_{s \in [0,T]}$ is progressively measurable for each $(y,z)$. Fix $(t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$ and let us denote by $(\{\varepsilon Y^{t,x,y,p}_s, \varepsilon Z^{t,x,y,p}_s\})_{s \in [0,t+\epsilon]}$ for $\varepsilon > 0$ small enough, the solution of the BSDE on $[0,t+\epsilon]$:

$$
\varepsilon Y^{t,x,y,p}_s = y + p \cdot (X^{t,x}_s - x) + \int_s^{t+\epsilon} g(u, \varepsilon Y^{t,x,y,p}_u, \varepsilon Z^{t,x,y,p}_u) \, du - \int_s^{t+\epsilon} \varepsilon Z^{t,x,y,p}_u \cdot dW_u. \quad (5)
$$

We claim the following proposition:

**Proposition 2.3** Let the assumptions (A 1), (A 2) and (A 4) hold for the function $g$ and let the notation (A 5) holds. Let us assume moreover that $\mathbb{E}\left[ \sup_{0 \leq u \leq T} |g(t,0,0)|^2 \right]$ is finite. Then, for each $(t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$, we have:

$$
L^2 = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \{ \varepsilon Y^{t,x,y,p}_t - y \} = g(t,y,\sigma^t(x)p) + p \cdot b(x).
$$
Proof. Since \((t, x, y, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n\) is fixed in this proof, we will drop, for notational convenience, the superscript \((t, x, y, p)\) and then write \((Y^\varepsilon, Z^\varepsilon)\) instead of \((Y^{t,x,y,p,\varepsilon}, Z^{t,x,y,p,\varepsilon})\).

Firstly, let us remark that by classical results on SDEs, see e.g. H. KUNITA [3], the terminal condition of the BSDE (5) is square integrable and then this BSDE has a unique solution in the space \(S^z_2 \times \mathcal{H}_2\).

Let us pick \(\varepsilon > 0\) and let us define, for \(t \leq s \leq t + \varepsilon\),

\[
\tilde{Y}^\varepsilon_s = Y^\varepsilon_s - (y + y \cdot (X^{t,x}_s - x)), \quad \text{and,} \quad \tilde{Z}^\varepsilon_s = Z^\varepsilon_s - \sigma^t(X^{t,x}_s)p.
\]

Using Itô’s formula, we easily show that, on the time interval \([t, t + \varepsilon]\), the process \((\tilde{Y}^\varepsilon, \tilde{Z}^\varepsilon)\) solves the following BSDE:

\[
\tilde{Y}^\varepsilon_t = \int_t^{t+\varepsilon} g(u, y + p \cdot (X^{t,x}_u - x) + \tilde{Y}^\varepsilon_u, \sigma^t(X^{t,x}_u)p + \tilde{Z}^\varepsilon_u) du + \int_t^{t+\varepsilon} p \cdot b(X^{t,x}_u) du - \int_t^{t+\varepsilon} \tilde{Z}^\varepsilon_u dW_u. \tag{6}
\]

Thus, by Proposition 2.2, we have the following estimate:

\[
\mathbb{E}\left\{ \sup_{t \leq s \leq t + \varepsilon} |\tilde{Y}^\varepsilon_s|^2 + \int_t^{t+\varepsilon} |\tilde{Z}^\varepsilon_u|^2 ds \middle| \mathcal{F}_t \right\} \leq C \mathbb{E}\left\{ 2^{2(K + K^2)} \varepsilon^2 \left( \int_t^{t+\varepsilon} |g(u, y + p \cdot (X^{t,x}_u - x), \sigma^t(X^{t,x}_u)p) + p \cdot b(X^{t,x}_u)| du \right)^2 \middle| \mathcal{F}_t \right\},
\]

and then, since \(g, b\) and \(\sigma\) are Lipschitz, we deduce that

\[
\mathbb{E}\left\{ \sup_{t \leq s \leq t + \varepsilon} |\tilde{Y}^\varepsilon_s|^2 + \int_t^{t+\varepsilon} |\tilde{Z}^\varepsilon_u|^2 ds \middle| \mathcal{F}_t \right\} \leq C_{x,p} \varepsilon^2 \mathbb{E}\left\{ 1 + \sup_{t \leq s \leq t + \varepsilon} \left( |X^{t,x}_s|^2 + |g(s, 0, 0)|^2 \right) \middle| \mathcal{F}_t \right\}.
\]

Taking the expectation in the previous inequality we obtain, since, by the additional assumption, \(\mathbb{E}\left[ \sup_{0 \leq s \leq T} (|X^{t,x}_s|^2 + |g(s, 0, 0)|^2) \right] \leq C(1 + |x|^2)\), the following estimate

\[
\mathbb{E}\left\{ \sup_{t \leq s \leq t + \varepsilon} |\tilde{Y}^\varepsilon_s|^2 + \int_t^{t+\varepsilon} |\tilde{Z}^\varepsilon_u|^2 ds \right\} \leq C \varepsilon^2, \tag{7}
\]

where \(C\) depends on \(x, y, p\) which is not important here since \((x, y, p)\) are fixed.

With this inequality in hands, it is easy to prove the result. Indeed, taking the conditional expectation in the BSDE (6), we get

\[
\frac{1}{\varepsilon} \{Y^{\varepsilon}_t - y\} = \frac{1}{\varepsilon} \tilde{Y}^\varepsilon_t = \frac{1}{\varepsilon} \mathbb{E}\left\{ \int_t^{t+\varepsilon} \left\{ g(u, y + p \cdot (X^{t,x}_u - x) + \tilde{Y}^\varepsilon_u, \sigma^t(X^{t,x}_u)p + \tilde{Z}^\varepsilon_u) + p \cdot b(X^{t,x}_u) \right\} du \middle| \mathcal{F}_t \right\}.
\]

We split the right hand side of the previous equality as follows:

\[
\frac{1}{\varepsilon} \{Y^{\varepsilon}_t - y\} = \frac{1}{\varepsilon} \mathbb{E}\left\{ \int_t^{t+\varepsilon} \left\{ g(u, y + p \cdot (X^{t,x}_u - x), \sigma^t(X^{t,x}_u)p) + p \cdot b(X^{t,x}_u) \right\} du \middle| \mathcal{F}_t \right\} + R_\varepsilon,
\]

where \(R_\varepsilon\) stands for

\[
\frac{1}{\varepsilon} \mathbb{E}\left\{ \int_t^{t+\varepsilon} \left\{ g(u, y + p \cdot (X^{t,x}_u - x) + \tilde{Y}^\varepsilon_u, \sigma^t(X^{t,x}_u)p + \tilde{Z}^\varepsilon_u) - g(u, y + p \cdot (X^{t,x}_u - x), \sigma^t(X^{t,x}_u)p) \right\} du \middle| \mathcal{F}_t \right\}.
\]
It is very easy to check that \( R_\varepsilon \) goes to 0 in \( L^2 \) when \( \varepsilon \) tends to \( 0^+ \). Indeed, since \( g \) is Lipschitz, we have
\[
|R_\varepsilon| \leq \frac{K}{\varepsilon} \mathbb{E}\left\{ \int_t^{t+\varepsilon} \left( |\tilde{Y}_u^\varepsilon| + |\tilde{Z}_u^\varepsilon| \right) du \right\},
\]
and then using Hölder’s inequality we obtain
\[
\mathbb{E}[|R_\varepsilon|^2] \leq \frac{K^2}{\varepsilon^2} \mathbb{E}\left[ \left( \int_t^{t+\varepsilon} \left( |\tilde{Y}_u^\varepsilon| + |\tilde{Z}_u^\varepsilon| \right) du \right)^2 \right] \leq \frac{2K^2}{\varepsilon} \mathbb{E}\left\{ \int_t^{t+\varepsilon} (|\tilde{Y}_u^\varepsilon|^2 + |\tilde{Z}_u^\varepsilon|^2) du \right\}.
\]
Taking into account the estimate (7), the previous inequality yields \( \mathbb{E}[|R_\varepsilon|^2] \leq C(\varepsilon^2 + \varepsilon) \) which shows the convergence of \( R_\varepsilon \) to 0.

It remains only to check that, as \( \varepsilon \to 0^+ \), in the sense of \( L^2 \),
\[
\frac{1}{\varepsilon} \mathbb{E}\left\{ \int_t^{t+\varepsilon} \left\{ g(u, y + p \cdot (X_{u}^{t,x} - x), \sigma^t(x)p) + p \cdot b(X_{u}^{t,x}) \right\} du \right\} \to g(t, y, \sigma^t(x)p) + p \cdot b(x).
\]
For this task, we set
\[
Q_\varepsilon = \frac{1}{\varepsilon} \mathbb{E}\left\{ \int_t^{t+\varepsilon} \left\{ g(u, y + p \cdot (X_{u}^{t,x} - x), \sigma^t(x)p) - g(u, y, \sigma^t(x)p) + p \cdot (b(X_{u}^{t,x}) - b(x)) \right\} du \right\},
\]
and we will prove that \( Q_\varepsilon \to 0 \) in \( L^2 \) and that the same is true for the remaining term, i.e.
\[
P_\varepsilon = \frac{1}{\varepsilon} \mathbb{E}\left\{ \int_t^{t+\varepsilon} \left\{ g(u, y, \sigma^t(x)p) - g(t, y, \sigma^t(x)p) \right\} du \right\}.
\]
For the first part, we remark that, since \( g, b \) and \( \sigma \) are Lipschitz,
\[
|Q_\varepsilon| \leq \frac{C}{\varepsilon} \mathbb{E}\left\{ \int_t^{t+\varepsilon} |X_{u}^{t,x} - x| du \right\},
\]
and then, by Hölder’s inequality, we get
\[
\mathbb{E}[|Q_\varepsilon|^2] \leq \frac{C^2}{\varepsilon^2} \mathbb{E}\left[ \left( \int_t^{t+\varepsilon} |X_{u}^{t,x} - x| du \right)^2 \right] \leq \frac{C^2}{\varepsilon} \mathbb{E}\left\{ \int_t^{t+\varepsilon} |X_{u}^{t,x} - x|^2 du \right\}.
\]
But since we have (see e.g. H. KUNITA [3]), for each \( r \geq 1 \),
\[
\mathbb{E}\left[ \sup_{0 \leq s \leq T} |X_{s}^{t,x}|^r \right] \leq C_T (1 + |x|^r),
\]
the function \( u \mapsto \mathbb{E}\left[ |X_{u}^{t,x} - x|^2 \right] \) is continuous. This function is equal to 0 at time \( t \), from which we deduce that \( Q_\varepsilon \to 0 \) in \( L^2 \).

Finally, we have, by Hölder’s inequality,
\[
\mathbb{E}[|P_\varepsilon|^2] \leq \mathbb{E}\left[ \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \left| g(u, y, \sigma^t(x)p) - g(t, y, \sigma^t(x)p) \right|^2 du \right].
\]
But by the assumption (A 4), \( \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \left| g(u, y, \sigma^t(x)p) - g(t, y, \sigma^t(x)p) \right|^2 du \) tends to 0 \( \mathbb{P}\)-a.s., and moreover
\[
\frac{1}{\varepsilon} \int_t^{t+\varepsilon} \left| g(u, y, \sigma^t(x)p) - g(t, y, \sigma^t(x)p) \right|^2 du \leq C\left( \sup_{0 \leq s \leq T} |g(s, 0, 0)|^2 + |y|^2 + ||\sigma^t(x)p||^2 \right)
\]
which is assumed to be integrable. Thus the result follows from Lebesgue’s theorem. The proof is complete. \qed

Remark. As we can see in the proof, the continuity of the process \((g(t, y, z))_{t \in [0, T]}\) (assumption (A 4)) is not really needed. We can prove the result if this process is only right-continuous.

Remark. It is worth noting that the assumption “\(\mathbb{E} \sup_{0 \leq t \leq T} |g(t, 0, 0)|^2\) finite” holds when \(g\) satisfies (A 3) or in the Markovian situation (see the subsection 3.2 below).

3 The deterministic case

3.1 Main result

This subsection is devoted to the study of the deterministic case for which we can prove a useful property of \(g\)-expectation. In this paragraph, \(g\) is defined from \([0, T] \times \mathbb{R} \times \mathbb{R}^d\) into \(\mathbb{R}\) and satisfies (A 1) and (A 3). Since \(\mathcal{E}_g\) is a generalization of the expectation, a natural question which arises is the following: if \(\xi\) is independent of \(\mathcal{F}_t\), do we have \(\mathcal{E}_g(\xi | \mathcal{F}_t) = \mathcal{E}_g(\xi)\)? We claim the following proposition which is mainly contained in [2].

**Proposition 3.1** If \(g\) is deterministic, then, for each \(\xi \in L^2(\mathcal{F}_T)\), we have \(\mathcal{E}_g(\xi | \mathcal{F}_t) = \mathcal{E}_g(\xi)\) as soon as \(\xi\) is independent of \(\mathcal{F}_t\).

**Proof.** It is enough to check, from the assumption (A 3), that \(\mathcal{E}_g(\xi | \mathcal{F}_t)\) is deterministic. Indeed, by construction we have \(\mathcal{E}_g(\xi) = \mathcal{E}_g(\xi | \mathcal{F}_t)\) and, under the assumption (A 3) (see S. Peng [8, Lemma 36.3]) for each constant \(c\), \(\mathcal{E}_g(c) = c\). We use the shift method to see that \(\mathcal{E}_g(\xi | \mathcal{F}_t)\) is deterministic. Let us introduce, for each \(s\) s.t. \(0 \leq s \leq T - t\), \(W'_{s} = W_{t+s} - W_{t}\). Then \(\{W'_s, 0 \leq s \leq T - t\}\) is a Brownian motion w.r.t. its filtration \(\mathcal{F}'_t\) which is nothing but \(\mathcal{F}'_{t+s}\) the \(\sigma\)-algebra generated by the increments of the Brownian motion \(W\) after the time \(t\). Since \(\xi\) is \(\mathcal{F}_T\)-measurable and independent of \(\mathcal{F}_t\), \(\xi\) is measurable w.r.t. \(\mathcal{F}'_{T-t}\). As a consequence, we can construct the solution \((Y_{s}', Z_{s}')_{0 \leq s \leq T-t}\) of the BSDE

\[
Y_{s}' = \xi + \int_s^{T-t} g(t + u, Y_{u}', Z_{u}')du - \int_s^{T-t} Z_{u}' \cdot dW_{u}', \quad 0 \leq s \leq T - t,
\]

and, setting \(s = v - t\), we get

\[
Y_{v-t}' = \xi + \int_{v-t}^{T-t} g(t + u, Y_{u}', Z_{u}')du - \int_{v-t}^{T-t} Z_{u}' \cdot dW_{u}', \quad t \leq v \leq T.
\]

Now, if we make the change of variables \(r = t + u\) in the integrals, we deduce that

\[
Y_{v-t}' = \xi + \int_v^{T} g(r, Y_{r-t}', Z_{r-t}')dr - \int_v^{T} Z_{r-t}' \cdot dW_{r-t}, \quad t \leq v \leq T,
\]

and

\[
Y_{v-t}' = \xi + \int_v^{T} g(r, Y_{r-t}', Z_{r-t}')dr - \int_v^{T} Z_{r-t}' \cdot dW_{r}, \quad t \leq v \leq T.
\]

It follows that \((Y_{v-t}', Z_{v-t}')_{v \in [t, T]}\) is a solution of the BSDE (1) on the time interval \([t, T]\) and then by uniqueness \((Y_{v}', Z_{v}')_{v \in [t, T]} = (Y_{v}, Z_{v})_{v \in [t, T]}\). In particular, \(Y_0' = Y_t\), and since \(Y_0'\) is deterministic, the proof is complete. \qed
Let us set $T$. To show that, let us not hold.

Indeed, for any $(t, y, z)$, we have, by Proposition 2.3,

$$g(t, y, z) = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \{ \mathcal{E}_g(y + z \cdot (W_{t+\varepsilon} - W_t) \mid \mathcal{F}_t) - y \} = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \{ \mathcal{E}_g(y + z \cdot (W_{t+\varepsilon} - W_t)) - y \},$$

and thus $g$ is deterministic.

We give a counterexample when $g$ is not deterministic. Actually, even in the simplest case, the linear case which corresponds to a Girsanov change of measures, the above property does not hold.

To show that, let us fix $T > 0$ and let us pick $t \in [0, T]$. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous and bounded function. We define, for each $(s, z) \in [0, T] \times \mathbb{R}$, $g(s, z) := f(W_{s \wedge t})z$. Moreover, let us set $\xi = W_T - W_t$ which of course is independent of $\mathcal{F}_t$. By classical results for linear BSDEs, see e.g. [2, Proposition 2.2],

$$\mathcal{E}_g(\xi \mid \mathcal{F}_t) = \mathbb{E}\left( \xi \exp \left\{ \int_t^T f(W_{s \wedge t})dW_s - \frac{1}{2} \int_t^T f^2(W_{s \wedge t})ds \right\} \mid \mathcal{F}_t \right) = \mathbb{E}\left( (W_T - W_t) \mid \mathcal{F}_t \right),$$

where $\mathbb{Q}$ is the probability measure on $(\Omega, \mathcal{F}_T)$ whose density w.r.t. $\mathbb{P}$ is given by

$$Z_T := \exp \left\{ \int_0^T f(W_{s \wedge t})dW_s - \frac{1}{2} \int_0^T f^2(W_{s \wedge t})ds \right\}.$$ 

But under the measure $\mathbb{Q}$, $\{ \tilde{W}_r = W_r - \int_0^r f(W_{s \wedge t})ds, 0 \leq r \leq T \}$ is a Brownian motion.

Moreover $W_T - W_t = \tilde{W}_T - \tilde{W}_t + \int_t^T f(W_{s \wedge t})ds = \tilde{W}_T - \tilde{W}_t + (T - t)f(W_t)$. It follows immediately that $\mathcal{E}_g(\xi \mid \mathcal{F}_t) = (T - t)f(W_t)$ and thus, if $f$ is not constant, we see that $\mathcal{E}_g(W_T - W_t \mid \mathcal{F}_t)$ is not deterministic which gives the desired result.

We are now able to answer the question asked in the introduction in the deterministic context.

Let, for $i = 1, 2$, $g_i : [0, T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$. We claim the following result:

**Theorem 3.2** Let the assumptions (A 1), (A 3) and (A 4) hold for $g_i$, $i = 1, 2$. Assume moreover that,

$$\forall \xi \in L^2(\mathcal{F}_T), \quad \mathcal{E}_{g_1}(\xi) \leq \mathcal{E}_{g_2}(\xi).$$

Then, we have,

$$\forall t \in [0, T], \forall (y, z) \in \mathbb{R} \times \mathbb{R}^d, \quad g_1(t, y, z) \leq g_2(t, y, z).$$

**Proof.** Let us fix $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$, and, for $n \in \mathbb{N}^*$ large enough, let us consider $\xi_n = y + z \cdot (W_{t+(1/n)} - W_t)$. Since the assumption (A 3) holds (see the remark after the proof of Proposition 2.3), we have, by Proposition 2.3,

$$n(i Y^n - Y) \to g_1(t, y, z), \quad as \ n \to \infty, \quad in \ L^2,$$

where $(i Y^n, i Z^n)$ is the solution of the BSDE

$$i Y^n_t = \xi_n + \int_t^T g_i(r, i Y^n_r, i Z^n_r)dr - \int_t^T i Z^n_r \cdot dW_r, \quad 0 \leq t \leq T.$$
In other words,

\[ n \{ \mathcal{E}_{g_i}(\xi_n \mid \mathcal{F}_t) - y \} \rightarrow g_i(t, y, z), \quad \text{in } L^2 \]

On the other hand, since \( \xi_n \) is independent of \( \mathcal{F}_t \), Proposition 3.1 ensures that, for \( i = 1, 2 \), \( \mathcal{E}_{g_i}(\xi_n \mid \mathcal{F}_t) = \mathcal{E}_{g_i}(\xi_n) \) and then by the hypothesis we deduce that

\[ n \{ \mathcal{E}_{g_i}(\xi_n \mid \mathcal{F}_i) - y \} \leq n \{ \mathcal{E}_{g_i}(\xi_n \mid \mathcal{F}_i) - y \} \cdot \]

Passing to the limit when \( n \) goes to infinity, we obtain, since \( g_1 \) and \( g_2 \) are deterministic, the inequality \( g_1(t, y, z) \leq g_2(t, y, z) \), which concludes the proof, since \( (t, y, z) \) is arbitrary and both functions \( g_1(\cdot, y, z) \) and \( g_2(\cdot, y, z) \) are continuous at the point \( T \).

\[ \square \]

Remark. As we can see in the proof, we only need to assume that \( \mathcal{E}_{g_i}(\xi) \leq \mathcal{E}_{g_i}(\xi) \) for \( \xi \) of the form \( y + z \cdot (W_s - W_t) \), for each \( (t, s, y, z) \), to get the result of this theorem. Actually, we can weaken a little bit more the assumption since it is enough to have the property when \( s - t \) is small enough, say less than \( \delta \), and this \( \delta \) may depend on \( (y, z) \).

### 3.2 An application to PDEs

We give in this subsection an application of the techniques described before to partial differential equations (PDEs for short). Semilinear PDEs constitute one of the field of applications of the theory of BSDEs as it was revealed by S. Peng [6] in the classical case and by E. Pardoux and S. Peng [5] for viscosity solutions of PDEs. Our setting is very close to that of F. Pradeilles [10].

Let \( f : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R} \) be a function s.t.

(A 6). There exist two constants \( K \geq 0 \) and \( q \geq 1 \), s.t.

1. \( \forall(y, z), \quad x \rightarrow f(x, y, z) \) is continuous;
2. \( \forall x, \forall(y, z), (y', z'), \quad |f(x, y, z) - f(x, y', z')| \leq K(|y - y'| + |z - z'|); \)
3. \( \forall x, \quad |f(x, 0, 0)| \leq K(1 + |x|^q). \)

In addition, let us consider \( h : \mathbb{R}^n \rightarrow \mathbb{R} \) a continuous function with polynomial growth. Then it is by now well known that BSDEs in Markovian context give a nonlinear Feynman-Kac formula for semilinear PDEs in the sense that the viscosity solution, say \( u \), of the semilinear PDE

\[ \partial_t u = \frac{1}{2} \operatorname{tr}(\sigma \sigma^t(x)D^2u) + b(x) \cdot \nabla u + f(x, u, \sigma^t(x)\nabla u), \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^n, \quad u(0, \cdot) = h(\cdot), \quad \text{(8)} \]

is equal, for each \( (t, x) \), to \( Y_{t,x} \) where \( \{(Y_{s,t,x}^s, Z_{s,t,x}^s)_{s \in [0, t]} \} \) is the solution of the BSDE

\[ Y_{s,t,x} = h(X_{t,x}^0) + \int_s^t f(X_{u,x}^0, Y_{u,x}^{t,x}, Z_{u,x}^{t,x})du - \int_s^t Z_{u,x}^{t,x} \cdot dW_u, \quad 0 \leq s \leq t, \]

where \( X_{t,x}^0 \) is the solution of the SDE (4).

Using this formula, we claim the following proposition:
Proposition 3.3 Assume that the assumption (A 6) holds for two functions \( f_1 \) and \( f_2 \) and that \( b \) and \( \sigma \) satisfies the assumption (A 5).

If, for \((x, y, p) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n\), there exists \( \eta > 0 \) s.t. \( \forall \varepsilon < \eta, u_1(\varepsilon, x) \leq u_2(\varepsilon, x) \), where \( u_i \) is the viscosity solution of the PDE (8) with the semilinear part \( f_i \) and the initial condition equal to \( h(\cdot) = y + p \cdot (\cdot - x) \), then \( f_1(x, y, \sigma^i(x)p) \leq f_2(x, y, \sigma^i(x)p) \).

Proof. Let us rewrite the assumption in a probabilistic way using the nonlinear Feynman-Kac formula mentioned above. Fix \((x, y, p) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n\). Then, there exists \( \varepsilon > 0 \), s.t.

\[
\mathbb{E}\left[ \sup_{0 \leq s \leq T} |X_{s}^{0, x} - x|^{2q} \right] \leq C_{T}(1 + |x|^{2q}),
\]

and then, by the growth assumption on \( f_i \), \( \mathbb{E}\left[ \sup_{0 \leq s \leq T} |f(X_{s}^{0, x}, 0, 0)|^{2} \right] \) is finite.

But, the assumption yields

\[
\forall \varepsilon < \eta, \quad \frac{1}{\varepsilon}\{1Y_{0}^{\varepsilon, x} - y\} \leq \frac{1}{\varepsilon}\{2Y_{0}^{\varepsilon, x} - y\},
\]

and, on the other hand, by Proposition 2.3, in the sense of \( L^2 \),

\[
\frac{1}{\varepsilon}\{iY_{0}^{\varepsilon, x} - y\} \longrightarrow f_i(x, y, \sigma^i(x)p) + p \cdot b(x), \quad \text{as } \varepsilon \to 0^+.
\]

Since moreover \( \frac{1}{\varepsilon}\{iY_{0}^{\varepsilon, x} - y\} \) is deterministic the above convergence holds in \( \mathbb{R} \) which gives the result. \( \square \)

Remark. We do not suppose in this result that \( f_i(x, y, 0) = 0 \).

4 Converse comparison theorem and application

4.1 A converse to the comparison theorem for BSDEs

Since the work of S. Peng [7], it is known that, under the usual assumptions, if \( \mathbb{P} \), a.s., the inequality \( g_1(t, y, z) \leq g_2(s, y, z) \) holds, then for each \( \xi \), we have

\[
\forall t \in [0, T], \quad Y_t^{1}(\xi) \leq Y_t^{2}(\xi),
\]

where \( ((Y_{s}^{1}(\xi), Z_{s}^{1}(\xi)))_{0 \leq s \leq T} \) is the solution of the BSDE:

\[
Y_t^{i}(\xi) = \xi + \int_t^T g_i(s, Y_s^{i}(\xi), Z_s^{i}(\xi))ds - \int_t^T Z_s^{i}(\xi) \cdot dW_s, \quad 0 \leq t \leq T.
\]

We prove in the next theorem a converse to this result with the additional assumption (A 3).
Theorem 4.1 Let the assumptions (A1), (A3) and (A4) hold for $g_i$, $i = 1, 2$. Assume moreover that,
\[ \forall \xi \in L^2(\mathcal{F}_T), \quad \forall t \in [0, T], \quad Y^1_t(\xi) \leq Y^2_t(\xi), \]
then, we have, $\mathbb{P}$–a.s.,
\[ \forall t \in [0, T], \forall (y, z) \in \mathbb{R} \times \mathbb{R}^d, \quad g_1(t, y, z) \leq g_2(t, y, z). \]

**Proof.** Let us fix $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$, and, for $n \in \mathbb{N}$ large enough, let us consider
\[ \xi_n = y + z \cdot (W_{t+1/n} - W_t). \]
We have by Proposition 2.3
\[ n(Y^1_t(\xi_n) - y) \rightarrow g_1(t, y, z), \quad \text{in } L^2. \]
On the other hand, by the hypothesis we deduce that
\[ n\{Y^1_t(\xi_n) - y\} \leq n\{Y^2_t(\xi_n) - y\}. \]
Extracting a subsequence to get the convergence $\mathbb{P}$–a.s. and then passing to the limit when $n$ goes to infinity, we obtain, the inequality $\mathbb{P}$–a.s., $g_1(t, y, z) \leq g_2(t, y, z)$. By continuity, we obtain finally that $\mathbb{P}$–a.s.,
\[ \forall (t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d, \quad g_1(t, y, z) \leq g_2(t, y, z). \]

The proof is complete. \( \square \)

**Remark.** We give in this remark the main lines of a different proof of this result. This approach is based on a “nonlinear decomposition theorem of Doob–Meyer’s type” due to S. Peng [9]. For a given $\xi$, the assumption of the theorem says that the process $Y^2(\xi)$ is a $g_1$–supermartingale and actually it can be seen that it is a $g_1$–supermartingale in the strong sense. Therefore, we can apply Theorem 3.3 in [9] to see that this process is also a $g_1$–supersolution. From this we deduce easily that, for each $\xi$ and for each $0 \leq t_0 < t \leq T$,
\[ \frac{1}{t-t_0} \int_{t_0}^t \mathbb{E}\left\{ g_1(u, Y^1_a(\xi), Z^2_a(\xi)) \right\} du \leq \frac{1}{t-t_0} \int_{t_0}^t \mathbb{E}\left\{ g_2(u, Y^2_a(\xi), Z^2_a(\xi)) \right\} du, \]
and choosing, as in Z. Chen [1], $\xi = X_T$, where, for a given $(t_0, y_0, z_0)$, $X$ is the solution of the SDE
\[ X_t = y_0 - \int_{t_0}^t g_2(u, X_u, z_0) du + \int_{t_0}^t z_0 \cdot dW_u, \]
we obtain, letting $t \rightarrow t_0^+$, $g_1(t_0, y_0, z_0) \leq g_2(t_0, y_0, z_0)$.

4.2 The $y$–independent case

Now, we turn to the case when the generator does not depend on the variable $y$; in this context we can answer the question asked in the introduction.

In this subsection, let $g : \Omega \times [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ be a function which satisfies the assumption (A1) and (A3). We recall the following lemma from S. Peng [8]:

**Lemma 4.2** For each pair of random variables $(\xi, \eta)$ in $L^2(\mathcal{F}_T) \times L^2(\mathcal{F}_t)$, we have, for $t \leq r \leq T$:
\[ \mathcal{E}_g(\xi - \eta \mid \mathcal{F}_r) = \mathcal{E}_g(\xi \mid \mathcal{F}_r) - \eta, \]
and, in particular, for each $0 \leq s \leq t \leq T$, $\mathcal{E}_g(\xi - \mathcal{E}_g(\xi \mid \mathcal{F}_t) \mid \mathcal{F}_s) = 0.$
Actually, we can give a characterization of generators independent of $y$.

**Lemma 4.3** Let us assume that the assumptions $(A\,1)$, $(A\,3)$ and $(A\,4)$ hold for $g$. Let us suppose moreover that

$$\forall \xi \in L^2(F_T), \forall x \in \mathbb{R}, \forall t \in [0,T], \quad \mathcal{E}_g(\xi - x \mid F_t) = \mathcal{E}_g(\xi \mid F_t) - x.$$ 

Then $g$ does not depend on $y$.

**Proof.** The proof is a direct application of Proposition 2.3. Indeed, if we pick a triple $(t, y, z)$ in $[0, T] \times \mathbb{R} \times \mathbb{R}^d$, we have,

$$g(t, y, z) = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \{ \mathcal{E}_g(y + z \cdot (W_{t+\varepsilon} - W_t) \mid F_t) - y \},$$

where the limit is taken in $L^2$. On the other hand, we have $\mathcal{E}_g(y + z \cdot (W_{t+\varepsilon} - W_t) \mid F_t) = y + \mathcal{E}_g(z \cdot (W_{t+\varepsilon} - W_t) \mid F_t)$ which yields

$$g(t, y, z) = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \mathcal{E}_g(z \cdot (W_{t+\varepsilon} - W_t) \mid F_t).$$

Hence $g$ does not depend on $y$. \hfill \Box

Let us consider, for $i = 1, 2$, $g_i : \Omega \times [0, T] \times \mathbb{R}^d \to \mathbb{R}$ s.t. $(g_i(s, z))_{s \in [0, T]}$ is progressively measurable for each $z$.

**Theorem 4.4** Let us assume that $(A\,1)$, $(A\,3)$ and $(A\,4)$ hold for $g_i$, $i = 1, 2$, and let us suppose that:

$$\forall \xi \in L^2(F_T), \quad \mathcal{E}_{g_1}(\xi) \leq \mathcal{E}_{g_2}(\xi),$$

then, $\mathbb{P} - a.s.$, we have

$$\forall s \in [0, T], \forall z \in \mathbb{R}^d, \quad g_1(s, z) \leq g_2(s, z).$$

**Proof.** Firstly, we show that under the assumptions of the theorem, for each $\xi \in L^2(F_T)$ and for each $t \in [0, T]$, $\mathcal{E}_{g_1}(\xi \mid F_t) \leq \mathcal{E}_{g_2}(\xi \mid F_t)$ and then we will apply the converse comparison theorem, Theorem 4.1 above, to get the result.

For each $\zeta \in L^2(F_T)$ and for each $t$, we have, by assumption,

$$\mathcal{E}_{g_1} \{ \zeta - \mathcal{E}_{g_1}(\zeta \mid F_t) \} \leq \mathcal{E}_{g_2} \{ \zeta - \mathcal{E}_{g_1}(\zeta \mid F_t) \},$$

and by Lemma 4.2, we get $\mathcal{E}_{g_1} \{ \zeta - \mathcal{E}_{g_1}(\zeta \mid F_t) \} = 0$ and

$$\mathcal{E}_{g_2} \{ \zeta - \mathcal{E}_{g_1}(\zeta \mid F_t) \} = \mathcal{E}_{g_2} \{ \mathcal{E}_{g_1}(\zeta \mid F_t) - \mathcal{E}_{g_1}(\zeta \mid F_t) \} = \mathcal{E}_{g_2} \{ \mathcal{E}_{g_1}(\zeta \mid F_t) - \mathcal{E}_{g_1}(\zeta \mid F_t) \},$$

and thus, we deduce that

$$\forall \zeta \in L^2(F_T), \forall t \in [0, T], \quad \mathcal{E}_{g_2} \{ \mathcal{E}_{g_1}(\zeta \mid F_t) - \mathcal{E}_{g_1}(\zeta \mid F_t) \} \geq 0.$$ 

Let us fix $\xi \in L^2(F_T)$ and $t \in [0, T]$. Applying the previous inequality to $\zeta = \xi \{ \eta < 0 \}$, where $\eta = \mathcal{E}_{g_2}(\xi \mid F_t) - \mathcal{E}_{g_1}(\xi \mid F_t)$, we get, noting that $\mathcal{E}_{g_1}(\xi \{ \eta < 0 \} \mid F_t) = 1_{\{\eta < 0\}} \mathcal{E}_{g_1}(\xi \mid F_t)$ since $\eta$ is $F_t$-measurable ([8, Lemma 36.6]), $\mathcal{E}_{g_2}(\eta 1_{\{\eta < 0\}}) \geq 0$. 

A converse comparison for BSDEs

113
On the other hand, since obviously the random variable \( \eta 1_{\{\eta < 0\}} \) is nonpositive, we obtain from the classical comparison theorem the fact that \( \mathcal{E}_{g_2}(\eta 1_{\{\eta < 0\}}) = 0 \). It follows that \( \mathcal{E}_{g_2}(\eta 1_{\{\eta < 0\}}) \leq \mathcal{E}_{g_2}(0) = 0 \). Thus, we can apply, in view of the assumption (A 3), Theorem 4.1, to obtain that, \( \mathbb{P} - a.s. \)

Thus, we can apply, in view of the assumption (A 3), Theorem 4.1, to obtain that, \( \mathbb{P} - a.s. \)

We will precise this result in a particular case but before we recall an elementary property of convex functions.

**Lemma 4.5** Let \( h : \mathbb{R}^n \rightarrow \mathbb{R} \) be a convex function and let us assume that \( h \) is bounded from above. Then \( h \) is constant.

**Proof.** Since \( h \) is convex, then, for each \( x \in \mathbb{R}^n \), \( h(\alpha y) \leq \alpha h(y) + (1 - \alpha)h(0) \) and then, for each \( \alpha \in [0, 1] \), choosing \( y = x/\alpha \) in the previous inequality, we get \( h(x) \leq h(x/\alpha) + (1 - \alpha)h(0) \). Since \( h \) is upper bounded, say by \( M \), the previous inequality yields \( h(x) \leq \alpha M + (1 - \alpha)h(0) \), and sending \( \alpha \) to 0+, we get \( \forall x \in \mathbb{R}^n, h(x) \leq h(0) \).

But, on the other hand, for each \( x \), \( 2h(0) \leq h(x) + h(-x) \) and by the previous inequality we deduce that \( 2h(0) \leq h(x) + h(0) \) which says that \( h(0) \leq h(x) \).

Thus \( h \) is constant. \( \square \)

**Remark.** As a byproduct of this lemma, we deduce that if \( g \) satisfies the assumptions (A 1) and (A 3) and if \( (y, z) \mapsto g(s, y, z) \) is convex then in fact \( g \) does not depend on \( y \). Indeed, if \( z \) is fixed, the function \( y \mapsto g(s, y, z) \) is convex. Moreover, since \( g(s, y, 0) = 0 \), the Lipschitz assumption gives \( |g(s, y, z)| \leq K\|z\| \) and thus by the previous lemma \( y \mapsto g(s, y, z) \) is constant.

Of course, the same result holds for concave functions.

This remark explains why we only consider functions of the variable \( z \) in the following result. Keeping the same setting as in the previous theorem, we claim the corollary:

**Corollary 4.6** Under the assumptions of Theorem 4.4, if we assume moreover that, \( \mathbb{P} - a.s. \), for each \( s, z \mapsto g_1(s, z) \) is convex (respectively \( z \mapsto g_2(s, z) \) is concave), then, if

\[
\forall \xi \in L^2(\mathcal{F}_T), \quad \mathcal{E}_{g_1}(\xi) \leq \mathcal{E}_{g_2}(\xi),
\]

there exists a progressively measurable bounded process \( (\alpha_t)_{t \in [0, T]} \) s.t. \( \mathbb{P} - a.s. \),

\[
\forall s \in [0, T], \forall z \in \mathbb{R}^d, \quad g_1(s, z) = g_2(s, z) = \alpha_s \cdot z.
\]

**Proof.** By Theorem 4.4 above, we deduce that

\[
\mathbb{P} - a.s. \quad g_1(s, z) \leq g_2(s, z).
\]

But the function \( z \mapsto (g_1 - g_2)(s, z) \) is convex and bounded from above by 0. Hence, we deduce, from Lemma 4.5, since \( g_1(s, 0) - g_2(s, 0) = 0 \), that \( g_1(s, z) = g_2(s, z) \). It follows that \( g_1(s, \cdot) \) is convex and concave and then \( g_1 \) is a linear map since \( g_1(s, 0) = 0 \). So there exists a progressively measurable process \( (\alpha_t)_{t \in [0, T]} \) s.t. \( \mathbb{P} - a.s. \),

\[
\forall s \in [0, T], \forall z \in \mathbb{R}^d, \quad g_1(s, z) = g_2(s, z) = \alpha_s \cdot z.
\]
Moreover, the process \( \alpha \) has to be bounded since the function \( g_1(s, z) \) is Lipschitz in \( z \) uniformly w.r.t \( s \) by the assumption (A 1).

**Remark.** This corollary covers the case of linear generators. However, it is worth noting that linear generators are a particular case of the following situation: assume that \( g_i \) is a function which is odd in the variable \( z \) and even in the variable \( y \). Then, it is very easy to check that \( \xi \mapsto \mathcal{E}_{g_i}(\xi) \) is also odd. Therefore, if we have, for each \( \xi \), \( \mathcal{E}_{g_1}(\xi) \leq \mathcal{E}_{g_2}(\xi) \), the inequality is actually an equality. Thus by Z. Chen’s result, we have \( g_1 \equiv g_2 \).

## 5 On Jensen’s inequality

This short section is devoted to Jensen’s inequality for \( g \)-expectation, but before we recall an elementary property of convex functions.

**Lemma 5.1** Let \( h : \mathbb{R} \to \mathbb{R} \) be a convex function. Then, for each \( x \in \mathbb{R} \), we have:

\[
\begin{align*}
    h(\alpha x) &\leq \alpha h(x) + (1 - \alpha)h(0) \quad \text{if } \alpha \in [0, 1], \\
    h(\alpha x) &\geq \alpha h(x) + (1 - \alpha)h(0) \quad \text{if } \alpha \in ]0, 1[.
\end{align*}
\]

**Proof.** Let us fix \( x \in \mathbb{R}^n \) and let us set, for \( \alpha \in \mathbb{R} \), \( f(\alpha) = h(\alpha x) - (\alpha h(x) + (1 - \alpha)h(0)) \).

Then, \( f \) is a convex function on \( \mathbb{R} \). Since \( f(0) = f(1) = 0 \), we immediately get the result. □

Now, we give a counterexample to the Jensen inequality. Consider \( g : \mathbb{R} \to \mathbb{R} \) defined as follows: \( g(z) = z^2 \) if \( z \in [-1, 1] \) and \( g(z) = 2|z| - 1 \) if \( |z| > 1 \). Let us introduce, for a fixed \( \sigma \in [0, 1], \xi := -\sigma^2 T + \sigma W_T \). Obviously, the solution of the BSDE

\[
Y_t = \xi + \int_t^T g(Z_s)ds - \int_t^T Z_sdW_s, \quad 0 \leq t \leq T,
\]

is \( ((-\sigma^2 t + \sigma W_t, \sigma))_{t \in [0, T]} \). Hence, for each \( t \in [0, T] \), we have:

\[
\frac{1}{2} \mathcal{E}_{g}(\xi | \mathcal{F}_t) = -\frac{\sigma^2}{2} t + \frac{\sigma}{2} W_t.
\]

Let us remark that, the solution of the BSDE

\[
Y_t = \frac{1}{2} \xi + \int_t^T g(Z_s)ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T,
\]

is given by \( \left\{ \left( \frac{\sigma}{2} W_t - \frac{\sigma^2}{2} T + \frac{\sigma}{2}(T-t), \frac{\sigma}{2} \right), \ t \in [0, T] \right\} \). It follows that:

\[
\frac{1}{2} \mathcal{E}_{g}(\xi | \mathcal{F}_t) - \mathcal{E}_{g}\left( \frac{1}{2} \xi | \mathcal{F}_t \right) = (T-t)\frac{\sigma^2}{4},
\]

which is positive as soon as \( t < T \).

This contradicts Jensen’s inequality in the simplest case i.e. the function is linear \( (x \mapsto x/2) \) and the generator is convex which implies, by the comparison theorem, that \( \xi \mapsto \mathcal{E}_{g}(\xi) \) is convex.

We end this paragraph by studying the Jensen inequality in a particular case: \( g \) independent of \( y \) and \( g \) convex in \( z \).
We assume moreover that \( P - a.s., \ z \rightarrow g(s, z) \) is convex for each \( s \), and we fix \( \xi \in L^2(F_T) \). Let us consider a function \( F : \mathbb{R} \rightarrow \mathbb{R} \) which is convex and s.t. the random variable \( F(\xi) \) is square integrable. By \( \partial F(x) \) we mean the subdifferential of the convex function \( F \) at the point \( x \) (see e.g. R. T. Rockafellar [11]) which, in our context, reduces to the interval \([F'_g(x), F'_g(x)]\).

**Proposition 5.2** If \( P - a.s. \), \( \partial F(\mathcal{E}_g(\xi \mid F_t)) \cap [0, 1[^c \neq \emptyset \), then

\[
\mathbb{P} - a.s. \quad F(\mathcal{E}_g(\xi \mid F_t)) \leq \mathcal{E}_g(F(\xi) \mid F_t).
\]

**Proof.** By the assumption on the subdifferential of \( F \), we have

\[
F(\xi) - F(\mathcal{E}_g(\xi \mid F_t)) \geq \beta(\xi - \mathcal{E}_g(\xi \mid F_t)),
\]

where \( \beta \) is an \( F_t \)-measurable random variable with values in \([0, 1[^c \). Let us pick a positive integer \( n \). Multiplying the previous inequality by \( 1_{\Omega_n} \) where \( \Omega_n \) stands for the set \( \{\mathcal{E}_g(\xi \mid F_t)| + |\beta| \leq n\} \) we have, setting moreover \( \beta_n = \beta_n \Omega_n \),

\[
1_{\Omega_n}[F(\xi) - F(\mathcal{E}_g(\xi \mid F_t))] \geq \beta_n(\xi - \mathcal{E}_g(\xi \mid F_t)).
\]

It is worth noting that the four terms of the previous inequality are square integrable since \( \xi \), \( F(\xi) \) belong to \( L^2(F_T) \), \( \beta_n \) is bounded by \( n \) and \( F \) is continuous on \( \mathbb{R} \) as a convex function. Thus, taking into account the fact that \( \mathcal{E}_g(\cdot \mid F_t) \) is nondecreasing, we deduce that

\[
\mathcal{E}_g(1_{\Omega_n}[F(\xi) - F(\mathcal{E}_g(\xi \mid F_t))] \mid F_t) \geq \mathcal{E}_g(\beta_n(\xi - \mathcal{E}_g(\xi \mid F_t)) \mid F_t)
\]

and since \( \beta_n \) and \( 1_{\Omega_n} \) are \( F_t \)-measurable, Lemma 4.2 implies

\[
\mathcal{E}_g(1_{\Omega_n}F(\xi) \mid F_t) - 1_{\Omega_n}F(\mathcal{E}_g(\xi \mid F_t)) \geq \mathcal{E}_g(\beta_n \xi \mid F_t) - \beta_n \mathcal{E}_g(\xi \mid F_t).
\]

On the other hand, since \( \beta_n \) is \( F_t \)-measurable and since \( \beta_n \in [0, 1[^c \), we have the inequality

\[
\mathcal{E}_g(\beta_n \xi \mid F_t) \geq \beta_n \mathcal{E}_g(\xi \mid F_t).
\]

Indeed, for \( t \leq r \leq T \),

\[
\beta_n Y_r = \beta_n \xi + \int_r^T \beta_n g(s, Z_s)ds - \int_r^T \beta_n Z_s \cdot dW_s,
\]

and since \( g \) is convex and \( \beta_n \in [0, 1[^c \), Lemma 5.1 shows that, using the fact that \( g(s, 0) = 0 \), for each \( t \leq r \leq T \),

\[
\beta_n Y_r \leq \beta_n \xi + \int_r^T g(s, \beta_n Z_s)ds - \int_r^T \beta_n Z_s \cdot dW_s.
\]

As a byproduct of this inequality, we get that \( \beta_n \mathcal{E}_g(\xi \mid F_t) \leq \mathcal{E}_g(\beta_n \xi \mid F_t) \), and coming back to the inequality (9), we obtain the inequality

\[
\forall n \in \mathbb{N}^+, \quad 1_{\Omega_n} F(\mathcal{E}_g(\xi \mid F_t)) \leq \mathcal{E}_g(1_{\Omega_n}F(\xi) \mid F_t).
\]

If \( n \rightarrow +\infty \), \( 1_{\Omega_n} F(\xi) \) converges to \( F(\xi) \) in \( L^2(F_T) \) and since \( \xi \rightarrow \mathcal{E}_g(\xi \mid F_t) \) is a continuous map from \( L^2(F_T) \) into \( L^2(F_t) \) (see e.g. S. Peng [8, Lemma 36.9]) we conclude the proof of the Jensen inequality in this context letting \( n \) tends to infinity.

**Remark.** It is also possible to obtain the Jensen inequality when the function \( g \) is concave instead of being convex. In this context, this requires essentially the fact that the convex function \( F \) satisfies \( 0 \leq F'(x) \leq 1 \) (e.g. \( F(x) = 1/\pi\{x\arctan x - 1/2 \ln(1 + x^2)\} + x/2\).
References


